DOCUMENT RESUME

ED 104 124 FL 006 306

AUTHOR	Zimmer, Karl E.
TITLE	Some General Observations about Nominal Compounds. Working Papers on Language Universals, No. 5.
INSTITUTION	Stanford Univ., Calif. Committee on Linguistics.
PUB DATE	May 71
NOTE	23p.
EDRS PRICE	MF-$0.76 HC-$1.58 PLUS POSTAGE
DESCRIPTORS	Deep Structure; English; German; *Language Universals; Language Usage; *Linguistic Theory; *Morphology (Languages); *Nominals; Semantics; *Structural Analysis; Surface Structure; Syntax; Transformation Theory (Language)
IDENTIFIERS	*Nominal Compounds

ABSTRACT

The paper begins with a discussion of several recently proposed analyses of nominal compounds in English. It is then suggested that the relations which may appropriately underlie nominal compounds of the type Noun + Noun can best be defined negatively, i.e. by listing those relations between two nouns which cannot underlie compounds rather than those which can. It is further argued that in order for compound formation to take place the relation between the nouns in question must be "appropriately classificatory." A brief examination of abstract compounds in English and German leads to the conclusion that there are probably fewer systematic restrictions valid across languages on the formation of abstract than of concrete nominal compounds. The appendix to the paper contains a brief examination of English nominal compounds with a primary + tertiary stress pattern (e.g., bookstore) and nominal phrases with a secondary + primary stress pattern (e.g., morning coffee). (Author)

ED104124

SOME GENERAL OBSERVATIONS ABOUT NOMINAL COMPOUNDS*

Karl E. Zimmer

U.S. DEPARTMENT OF HEALTH, EDUCATION & WELFARE
NATIONAL INSTITUTE OF EDUCATION
THIS DOCUMENT HAS BEEN REPRODUCED EXACTLY AS RECEIVED FROM THE PERSON OR ORGANIZATION ORIGINATING IT. POINTS OF VIEW OR OPINIONS STATED DO NOT NECESSARILY REPRESENT OFFICIAL NATIONAL INSTITUTE OF EDUCATION POSITION OR POLICY

ABSTRACT

The paper begins with a discussion of several recently proposed analyses of nominal compounds in ENGLISH. It is then suggested that the relations which may appropriately underlie nominal compounds of the type Noun + Noun can best be defined negatively, i.e. by listing those relations between two nouns which cannot underlie compounds rather than those which can. It is further argued that in order for compound formation to take place the relation between the nouns in question must be "appropriately classificatory". A brief examination of abstract compounds in ENGLISH and GERMAN leads to the conclusion that there are probably fewer systematic restrictions valid across languages on the formation of abstract than of concrete nominal compounds. The appendix to the paper contains a brief examination of ENGLISH nominal compounds with a primary + tertiary stress pattern (e.g. bóokstòre) and nominal phrases with a secondary + primary stress pattern (e.g. mórning cóffee).

FL006306

---

* The present paper is intended as the 'theoretical' part of a study of nominal compounds of the structure Noun + Noun, i.e. compounds which denote a subset of the referents of the head noun (all bird dogs are dogs; none of them are birds). No particular attention will be paid to exocentric compounds. Examples to illustrate the points made here will be drawn principally from ENGLISH. In the second part of the study nominal compounds in a number of languages will be examined in the light of the considerations set forth below.

Considering the amount of effort that has over the years been devoted to the analysis of nominal compounds in various languages it seems quite remarkable that there is any aspect of these constructions which has not as yet been fully explored. I hope to show that such aspects do, however, exist, and that some recent treatments of compounds of considerable formal sophistication have failed to come to grips with certain very basic issues.

Before we ourselves come to grips with these issues some other points must be dealt with briefly. One of these is the productivity of compounding patterns. This matter has been dealt with by Lees (1960) and a number of subsequent writers,[1] one quotation from a recent treatment will do here: "A more fruitful approach to the entire problem is to ask whether nominal compounds ... are merely a list of items like a set of vocabulary or whether the languages involved possess certain mechanisms which allow the native speaker to use compounds creatively. If the former is true, there is no need for any underlying structure of compounds. One simply generates the compounds by listing them. If the latter is true, the task of the linguist is to describe such mechanisms." (Li 1970, 68). Clearly it is the latter which is true for many types of compounds. It is not necessary for a speaker of ENGLISH to have previously encountered and memorized, as a unit, the compound mustard stain in order to ask the dry cleaner to please be sure and get that mustard stain out. This leaves us with the problem of deciding just what the nature of the compound-generating mechanism ought to be.

In the recent literature we find two types of proposals for such mechanisms. What can be described as a non-semantic mechanism is presented in considerable detail by Lees (1960). Botha (1968), in his study of AFRIKAANS nominal compounds, uses a basically similar model, although he notes certain inadequacies and proposes some additions which we need not go into here. Basically this model generates nominal compounds by means of transformational rules from the same kinds of deep structures that underlie full sentences and the deep structures that the compounds are derived from are themselves generated by rewrite rules which refer only to syntactic categories and features. In other words, the approach to syntax involved here is, for the earlier version (e.g. Lees 1960), that of Chomsky (1957), and, for the later version (e.g. Botha 1968) that of Chomsky (1965).[2] To give a concrete example, marrow bone

---

[1] E.g. Schachter 1962, Householder 1962, Botha 1968.

[2] I shall on the whole ignore aspects of the non-semantic model which an up-to-date proponent of it would by now have abandoned because of fundamental changes in the area of syntactic theory. The general context in which it will be discussed is thus that of the Chomskyan Aspects view of syntax and semantics. For a discussion by Lees of compound generation within this general framework see Lees (1966).

is derived from The bone has marrow (or rather from the deep structure underlying that sentence; henceforth this provision will be taken for granted) via bone which has marrow and bone with marrow (Lees 1960, 133). It is of course clear that some of the transformations that have been assumed to be operative in the derivation of compounds must be of a very special type not encountered elsewhere in grammars. If dragon poison could be derived, among other things, both from The poinson is from a dragon and The poison is for a dragon (cf. Lees 1960, 122-3) then some elements which are relevant to the semantic interpretation of these sentences have been deleted irrecoverably. This point was first raised in a review of Lees' book by Rohrer (1966), and has recently been discussed by Lees (1970).[3] Lees notes that in trying to account for how a grammatical analysis with verb-deleting transformations might underlie a hearer's interpretation of an arbitrary compound of the subject-object type (e.g. bedbug from The bug infests beds) we must assume either (1) that the hearer understands every such compound to be as many ways ambiguous as there are different verbs which could have appeared in the underlying structure or (2) that the grammar associates explicitly with each object-and-subject pair just those verbs which may be deleted in the formation of allowable compounds (Lees 1970, 180). Lees then suggests an analysis of compounds which, in contrast with the one he had proposed earlier, is no longer non-semantic. We shall return to this proposal below.

In a new study by Lila and Henry Gleitman (1970) syntactic rules for the generation of compounds are given--the great majority of compounds being derived from relative clauses, e.g. mail-dog from the dog who brings the mail--but some questions are raised about a transformational derivation of compounds à la Lees (1960). "There are two problems," the Gleitmans observe, "that make it difficult to treat compounds as transformational processes: there are systematic difficulties in recovering deleted material, and the association between relatives and compounds is not symmetrical, not quite meaning-preserving.[4] At the same time, what is deleted is not utterly unknown (good guesses can be made at the appropriate verb), and there is a quasi-paraphrastic, close to meaning-preserving, relation. We have reflected the systematic,

---

[3] See now also Gleitman (1970, 91-4).

[4] It has been noted earlier that compounds are generic ("... not every man who removes the garbage is a garbage-man. Only a man who occupationally, customarily, eternally removes the garbage is a garbage-man. ...man who removes the garbage is a definition of garbage-man, but the converse is not sensible.") (96).

generative, nature of the relation in rule [iv] [the rule that generates e.g. mail-dog], and we have claimed as well (by specifying proVB) that the underlying verb is not recoverable by any generative process that we know of. We have thus implicitly relegated the problem of the appropriate verb to the semantic component of the grammar, where we have much company in our ignorance." (Gleitman 1970, 96-7). They conclude that it is to be regarded as unlikely "that compounds and relative clauses will be found to be in a derivational relation." and express their belief "that this process of word formation will appear as a systematic part of the lexicon, along with the various affixing rules ... which it closely resembles." (97). The Gleitman model can be regarded as primarily syntactic with semantic question marks.

The other type of mechanism for compound generation we can call semantic. Proposals for such mechanisms can be found in Brekle (1970), which presents the most carefully worked out model, Lees (1970), and Li (1970). Li's analysis is as yet very preliminary, and clearly inadequate in some important respects. A detailed discussion of it would be premature; suffice it to say that Li proposes a convention which allows two nouns to be concatenated to form a compound noun, and that each one of his compound-generating mechanisms consists of this convention together with a condition specifying the semantic relationship between the constituents of the compounds it generates. Thus the semantic relation for one type of compound is specified as 'N_2 denotes a protective device against N_1', and the compounding mechanism with which this semantic specification is associated has as its output compounds such as sunglasses, gas mask, flea collar, fly paper, raincoat (cf. Li 1970, 93-111).

Brekle (1970) develops his system on the basis of semantic deep structures represented in terms of the calculus of functions of symbolic logic. These deep structures represent sentence contents ('Satzbegriffe') from which nominal compounds are derived by rules of several different types. One of the most important features of Brekle's system are his rules of topicalization, which enable him to derive compounds with different head constituents from the same sentence content. Thus from a sentence content which can informally be represented as 'some fish being in some pond' it would be possible to generate, by topicalizing fish, the compound pond fish, while the topicalization of pond would lead to fish pond. A detailed discussion of Brekle's proposals would take us considerably beyond the scope of the present paper. His study, while it deals with many important problems and contains a number of valuable insights, strikes me as having an insufficiently semantic basis for a model with a purportedly "satzsemantisch" deep structure. A brief look at some of Brekle's relational predicates will illustrate what I have in mind. Thus the symbol

CAUS represents a relation which is glossed as 'verursachen', and we are told: "Im Vorbereich dieser Relation treten [Prädikate] auf, die weiter durch das Merkmal "+ agens" spezifiziert werden müssen, d.h. als "Verursacher" von Ereignissen und Handlungen kommt nur eine bestimmte Untermenge aller [Prädikate] sinnvollerweise infrage." (Brekle 1970, 116). The symbol AEFF represents either the relation 'affizieren' or the relation 'effizieren', and appears in deep structures as a complex term of the relation represented by CAUS (Brekle 1970, 117). When we look at compound types, we find that a deep structure of the form

CAUS [w, AEFF(R, y)]

is said to underlie bug spray, chastity belt, cowboy and tone arm, among other compounds (with topicalization of w and deletion of R)[5] (Brekle 1970, 163). This means that spray, belt, boy and arm are all regarded by Brekle as being characterized by the feature "+agens". It seems clear, however, that in any even minimally adequate semantic system only boy can be viewed as an agent in the cited compounds, while spray, belt and arm are all instruments. The same deep structure is furthermore postulated for Nixen hater (p. 164), with the symbol w being replaced by the agentive suffix -er. Thus the hater is regarded as a "Verursacher" and the hatee as an affected object, which seems, semantically speaking, at best a doubtful analysis. Similar examples could be cited in considerable numbers. In general, Brekle's purportedly semantic analysis seems to be far too much influenced by a traditional subject-and-object view of sentence structure; as a result of this surfacist bias the book is frequently marred by a facile identification of "agent" with "subject" and other errors of a similar kind. This is not to say that a compound-generating model of the general type proposed by Brekle is in principle incapable of providing a satisfactory analysis of compounding, but in its current state it is too much of a syntacto-semantic hybrid to be very persuasive.

We now return to an examination of the recent suggestions put forward by Lees (1970). Lees proposes that the deep structures underlying nominal compounds are best regarded as being of the type presented by Fillmore (1968), noting that "such an analysis permits a much finer distinction and a closer connection between the meanings of a compound and its deep syntactic structures," and that "it may also afford a decision on the problem of indiscriminate verb ellipsis." (Lees 1970, 181). Let us look at a couple of analyses in this framework: air rifle, motor car, steamboat, water wheel, windmill are all said to be derived from

---

[5] As an illustration, in the deep structure underlying cowboy w = boy, R = herd or some similar verb, y = cow.

a deep structure of the form

Verb--Object--Instrument

via an intermediate stage

Instrument V-s Object.[6]

Lees notes that the V of each member of this set of compound "may be viewed as a variant of the verbs impel, energize, activate, power, propel, etc." He then goes on to observe that "it may be possible to associate one, or a small number of generalized verbs with certain classes of compounds by fixed grammatical rule, so that the compounds in question need not be described by the grammar in such a way as to imply that they are indefinitely ambiguous." (Lees 1970, 181-2). By "generalized verb" is meant the minimal set of semantic features which characterize all the members of a set of verbs such as impel, energize, and so forth. Another group of compounds -- e.g. battle fatigue, blood stain, diaper rash, finger print, hay fever, saw dust, soap suds -- is analyzed as being derived from the same V-O-I deep structure, with the generalized verb in these cases having the semantic features common to cause, yield, engender, emit, produce, etc.[7] Some compound types are said not to

---

[6] A somewhat similar approach is outlined in Dokulil (1964), except that Dokulil, unlike Fillmore and Lees, regards the categories involved in word formation as purely conceptual and outside the province of syntax.

[7] A similar suggestion can be found in Gleitman (1970): "Nevertheless, in some way (unknown to us) a partial semantic specification does seem to be associated with lion-house. All of the relative clauses we gave as examples above [a house for a lion, a house belonging to lions, a house suitable for lions] had the sense of dwell, occupy. If one conceives a grammar in which only such general semantic features, and not specific verbs such as dwell, live, are the elements of base structure, then the objection we have raised [relating to different but compatible readings of a compound, e.g. the different paraphrases for lion-house just given] is effectively countered, and our decision to treat the verb ... as unspecified again seems questionable... More important, there are innumerable cases where compounds are amenable to many incompatible, semantically distinct interpretation... Here ambiguity, surely not vagueness, seems the better descriptive term. A rule ... which specifically calls for a pro-Verb does not explicate this ambiguity." (96-7).

require special assumptions about the underlying verb, since this verb can be regarded as determined by the meaning of the head noun (e.g. brick mason, gunsmith, horse doctor). After setting up seven main classes of compounds based on deep structures in which nouns appear as case-marked complements of verbs, Lees goes on to observe that "there are...still other large classes of compounds which may not reflect the syntactic relations of noun complements of a verb but rather those of the constituents of certain copula sentences, or those of the genitive constructions, etc." (Lees 1970, 185). He then gives a few examples of types "not yet fully investigated," such as Object/Property (collar size, vapor pressure), Whole/Part (arrowhead, cart wheel, whale bone), Form (brick cheese, fireball). His final conclusion is that "more sophisticated analysis in the study of compounding provides some evidence for the view that the deepest syntactic structure of expressions is itself a more or less direct picture of their semantic descriptions." (Lees 1970, 185).

It seems clear that Lees' new analysis is in some important respects superior to that in Lees (1960). For one thing, it is no longer necessary to worry about such pseudo-problems as whether battle fatigue should be derived from an underlying The battle causes fatigue or from The fatigue comes from battle. The deep structure underlying the compound would, in the new view, simply contain battle and fatigue as case-marked nouns (Lees suggests Instrument and Object, respectively); presumably these nouns have the same case functions in the two sentences cited above as possible non-semantic sources for battle fatigue.[8] Another advantage is that a solution is now proposed for the problem of indefinite ambiguity caused by verb deletion. It should be noted, however, that this is done at the expense of adding a new kind of lexical entry to the dictionary, namely the above mentioned "generalized verbs". These would apparently occur only in deep structures underlying nominal compounds, and would all be marked for deletion by a special deletion rule.

One objection that had been raised to Lees' earlier analysis still applies. This is that the semantic content of all compounds appears to be treated as being essentially similar, and given by the deep structures assumed

---

[8] There is of course the further problem of deciding whether the case functions proposed by Fillmore constitute an adequate universal set. Thus battle clearly differs in function from an instrument manipulated by an agent; it is not clear, however, whether it should nevertheless be regarded as Instrument, with the difference being ascribed to its inherent meaning, or whether something like a Cause case should be added to the inventory.

to underlie the compounds. But while such a treatment will do for compounds like car wheel, it will not do for lexicalised compounds like bedbug. A bedbug is not any bug that happens to infest a bed, but a particular kind of bug, and it is a bedbug whether it infests a bed or not. Lees' analysis may account for the original coining of the compound, but it does not account for its use as what might be called a complex lexical item. Clearly their tendency to become lexicalised is an important aspect of compounds, and should not be totally ignored in any treatment of this subject.[9]

In spite of some shortcomings, however, Lees' approach embodies an important insight, namely, that some sort of specification must be given for the functions that characterize the nouns entering into compounds. I had myself started out, when I first began to worry about compounds several years ago, with the pleasingly simple and general notion that any relation between the two nouns of an N+N compound would do.[10] In this view, the perception of any relation whatsoever between two nouns would provide a sufficient criterion for the generation of a compound, and all that a hearer could expect upon hearing some novel compound would be that the speaker using it had in mind "some sort of relation" between the constituent nouns. But this will not do, as can be seen by looking at a few ill-formed compounds. Thus we cannot get a compound bone dog meaning 'dog lacking a bone'.[11] Nor,

---

[9] To speak of a 'tendency to become lexicalised' and to leave it at that is a serious oversimplification of a complex problem. It seems fairly clear that some compounds are lexicalised as soon as they are created, or rather they are created as names for a particular referent. Thus I suspect that bedbug was probably never used to refer to any and all bugs that might be found in beds, but was rather coined as a name for cimex lectularius. Notice also such compounds as bulldog, grapefruit, dollarfish.

[10] This turns out not to be novel notion. Some sixty-five years ago Henry Bradley stated: "The general meaning of this class of compounds might be expressed by saying that the noun which is formed of the two nouns A and B means 'a B which has some sort of relation to an A or to A's in general." (Bradley 1906, 113).

[11] Motsch notes (apparently citing a conclusion reached by Heidolph in his dissertation): "Als semantisches Merkmal der Komposita ergibt sich z.B., daß sie keine privativen Beziehungen ausdrücken können. Mit diesem Merkmal stehen sie im Gegensatz zu Adjektivderivationen, die durch -los, -frei solche Verhältnisse bezeichnen können." (Motsch 1962, 49-50). Unfortunately Heidolph's dissertation (Beziehungen zwischen Kompositum und attributiven Substantivkonstruktionen in der deutschen Gegenwartssprache (Berlin, 1961)), was not available to me.

as John Ross has pointed out, can the relation 'A is between two B's' underlie a compound; house tree cannot mean 'tree standing between two houses'. Chocolate girl would not be generated or understood as meaning 'girl who hates chocolate'. In fact, one can go on indefinitely making up compounds which cannot be related to certain sentences or semantic structures in which the two constituent nouns appear. What is not entirely clear to me is whether 'positive' characterizations of compound types such as those proposed by Lees, Brekle, and Li (and, of course, many others) can ever achieve the goal of an exhaustive definition of acceptable compounds. By positive characterization I mean a pairing of surface compounds with some sort of underlying structure, with the ultimate goal that any acceptable compound must conform to one of the listed pairings. My impression in looking at most analyses of compounds is that some sort of generative apparatus is set up for attested compounds and that the analyst, having done this with whatever degree of exhaustiveness his memory and industry, as well as available dictionaries and texts, permit, assumes at the end of his endeavors that he has given a reasonably adequate characterization of what constitutes an acceptable compound in the language. An interesting departure from the procedure just outlined is made by Li, who proposes that there are compounds which should not be regarded as the product of any compounding mechanism that is to be imputed to speakers as part of their linguistic competence. Some examples are: cradle song 'a song to lull a child in the cradle to sleep', dishwater 'water in which one has washed dishes', death mask 'a cast of the face taken just after death', timber line 'the upper limit of tree growth on mountains and in the arctic region'. Li closes his discussion of these compounds as follows: "These compounds are not idioms as in the case of redcap, snakepit, etc. They are created almost explicitly for the sake of convenience. Each one is different from the other, and none of them form a type which may have a number of compounds. In other words, there is no productivity to speak of, and there exists no mechanism which generates them. The only way to generate them is by the trivial method of listing. In view of these compounds, one may conclude that compounds, in isolated instances, also serve as a means for telegraphic speech and [as a] construction [for] providing special terminologies or names." (Li 1970, 73-4). Li thus takes the position that there is a listable set of compounding patterns (he sets up twenty-four, which he finds implemented in both ENGLISH and CHINESE), and that these can be regarded as adequately delimiting the competence aspect of compound generation. There will presumably in all languages be some compounds which do not conform to these patterns, but these, Li holds, can appropriately be regarded as unpredictable and idiosyncratic, i.e. as lying outside the domain for which a systematic description is responsible. Although this view is attractive, I do not believe that it is correct. For one thing,

there are of course, other, similar compounds in existence (e.g. life mask, snow line). This, however, is a trivial objection, since the existence of two or three forms, rather than one, obviously is not sufficient to demonstrate the productivity of a pattern (thus suffixation with the diminutive -ling can hardly be regarded as a productive process in ENGLISH). Perhaps more serious is the objection that any existing compound can presumably serve as a model for the production of others, so that given dishwater someone might come up with footwater 'water in which one has washed one's feet'. This observation too, however, does not really get to the heart of the matter, which is whether the generation by a speaker of a compound with no apparent model should necessarily be regarded as somehow deviant. It is my feeling that it should not. Unfortunately it is impossible to prove conclusively that my position is the correct one; this is something everyone must decide for himself in solitary confrontation with his linguistic intuition. The best I can do is to provide some examples of unusual compounds and let the reader decide whether they unduly stretch the limits of ENGLISH. Let us assume (prompted by the suggestion of cradle song) that the sirens encountered by Odysseus had been native speakers of ENGLISH; it is my position that they could quite appropriately have referred to their song as ship song (which might be paraphrased as 'song sung to sailors in passing ships', or something of that sort), or death song ('song sung to lure sailors to their death'), or reef song ('song sung to lure sailors onto the reef'). Another example, this one with a modern context, is pumpkin bus. This was uttered on a bus which was transporting those members of an excursion group who did not want to stop and buy pumpkins on the way to Berkeley (the trip took place a few days before Halloween). As the bus was approaching a pumpkin field along the road it slowed down, whereupon someone exclaimed, "Hey, don't stop, this isn't a pumpkin bus!" If one wanted to paraphrase pumpkin bus one would have to do so by some expression such as 'bus for people who want to buy pumpkins', although it should be pointed out that the coiner of this compound did not of course have to have this or any other specific underlying structure in mind when he uttered it. I urge the reader to make up other unusual compounds for himself and to examine how normal or deviant they seem linguistically.

What the approach based on a positive characterization of compounds seems to imply is that, in order to produce a nondeviant compound, a speaker must in some way check its underlying structure against his list of possible compounding rules to see if it matches the input to one of them. I doubt very much that this is in any sense an accurate model of compound formation. Thus the fairly comprehensive list of compounds of the type

$$V - \underset{N_1}{Obj} - Ag - \underset{N_2}{Instr} \rightarrow \quad Ag\ V\text{-}s\ N_1\ with\ N_2 \rightarrow \quad N_1 + N_2$$

given by Lees (1970, 183) contains no subtype which would accomodate water rod, which seems to be an entirely unexceptionable compound in the sense of 'divining instrument' (unexceptionable in principle, that is; its actual formation with that meaning would, given the present lexical circumstances of ENGLISH, presumably be inhibited by the existence of the term divining rod). This compound could be paraphrased by something like 'Agent uses the rod to find water', and find is a verb which does not fit very well into any of the "generalized verb" niches provided by Lees. And of course the verb in this case is not reconstructible from the meaning of the head noun. Yet all this does not appear to make its generation deviant.[12] It may of course be true that in interpreting a compound a hearer would fit it into one of the more prevalent types, but it seems to be typically the case that when he does not succeed in doing so he will resort to more fanciful interpretations, or--if he has been successful--will nevertheless accept other, unexpected interpretations as long as they are not excluded by some general constraint (see below), even though he may not have had a model for such an interpretation in his repertoire of familiar compounds.

If the position just outlined is correct, where does this leave our earlier contention that not all conceivable relations between two nouns can appropriately underlie a compound? It is my impression that the most promising approach to reconciling these two at first glance apparently contradictory views is to abandon the attempt at an exhaustive positive characterization of compound types, and to pursue the approach suggested by the Heidolph-Motsch observation cited above, i.e. to try and define those relations which cannot underlie compounds. If that can be adequately done, one need then only say that any other relation can appropriately lead to the formation of a compound, even though it may not belong to one of the types commonly encountered in the language. These most common types, incidentally, seem to me to result from the frequent reference to certain characteristics which are apparently universally used for distinguishing different species of a genus (using the terms loosely) in the real world, e.g. appearance (sailfish), purpose (teaspoon), typical location (field mouse). To come back to what we might call a negative definition of acceptable compounds: whether this is a workable approach is a question that cannot be answered without a great deal of further study. All I can do at present is throw out a few hints. For one thing, the Heidolph-Motsch criterion seems essentially correct; knife box cannot conceivably be used to refer to a

[12] It is perhaps worth noting that water rod does not conform to any of the twenty-four compound types of Li (1970) either.

box which typically has no knives in it. As for location in space and time, coincidence seems to be required, i.e. location of the referent of one noun at or within the area or period referred to by the other, except that for space adjacence will apparently also do. Sea town is possible for a town on the coast, but not for a town far from the sea; river road could refer to a road that runs near a river as well as to a road leading to a river; Wednesday lecture cannot refer to a lecture given on Tuesday, Wednesday, or Thursday (a moment's reflection will show that such sentences as 'This week I'm giving my Wednesday lecture on Tuesday' are not real counterexamples to this claim). Relations which imply some sort of rejection are not appropriately mentioned in the compound. War hater is all right, but war man in the sense of 'man who dislikes, denounces, etc. war' is not (although it will do in the sense of 'man who is in favor of war'). One might perhaps say that any sort of distancing relationship is unlikely to be appropriate as the source of compounds, but "distancing relationship" is obviously such a vague term as to be nearly useless. That is, given some specific relation, it seems fairly unlikely that different judges would necessarily agree on whether it should or should not be classified as a distancing relation. And in this rather unsatisfactory state we shall, unfortunately, have to abandon the study of the relations which can, or cannot, underlie acceptable compounds.[13]

Having thus cleared the decks (not entirely, it is to be hoped, by sweeping things under the rug) we can proceed to the basic issues announced in the first paragraph of this paper. What I want to focus on are the necessary conditions that must be fulfilled for compounding to occur appropriately, given that some relationship obtains between two nouns that is not marked [-compounding].[14] Many discussions of compounds make the explicit or implicit assumption that compounding can be regarded as an optional process, i.e. that given a semantic (or syntactic, or semantactic) underlying structure of an appropriate form speakers are free either to derive compounds from it or to encode it in sentential form. (This statement does not apply to Li's compound-generating mechanism, which is distinct

---

[13] The usual distinction between occurring and grammatical forms of course applies here too. That is, the kinds of restrictions on compounding that we have discussed need not necessarily be abandoned if some compounds turn up that violate them -- as long, at least, as such apparent counterexamples can reasonably be regarded as somehow deviant.

[14] As a consequence of this emphasis it is non-lexicalised compounds which will receive most of our attention.

from the set of rules that produce sentences, nor to Brekle's, although in his system the only factor that consistently distinguishes the structures underlying compounds from those underlying sentences is the absence of the feature of assertion in the former.) This approach would imply that it is quite normal for someone to say Look at that street dog interchangeably with Look at that dog in the street, or Don't break the table glass interchangeably with Don't break the glass on the table. Since these expressions are in fact not interchangeable -- the sentences with compounds are surely both somewhat odd -- something more must be involved. Brekle has briefly examined what this something might be, confining himself, however, to the distinctive element in compounds of the type Adj + Noun (e.g. blackbird, madman). He suggests that the predicate 'habitually' or 'specifically' is involved (Brekle 1970, 148) ("der psychische Zustand eines mádmàn muß als HABITUELL gestört betrachtet werden, im Vergleich mit ähnlichen temporär begrenzten Zuständen der mád mén" [Brekle 1966, 21]).[15] This additional predicate is necessary in order to give expression to the fact that the set containing the referents of the compound is a subset of the set containing the referents of the syntactic group (all blackbirds are black birds, but there are black birds which are not blackbirds). The Gleitmans (1970) come back to this issue several times and at some length. For Adj + Noun compounds they note that "this pattern is used most often (though not exclusively) to indicate that a proper name -- some particular species or type -- is intended," (Gleitman 1970, 86) and then go on: "... the compound, by being a compound, implies a name, some unitary character to the relationship among the elements. In some sense, the outputs of this constructionally regular process tend to lose their constructional character... This problem exists similarly for various kinds of compounds... The compound becomes a name..." (87). They suggest a rule which can be paraphrased as

> An expression N_1 which is named $\underset{N}{Adj} + N_1$, with primary stress on the second element, becomes $\underset{N}{Adj} + N_1$, with primary stress on the first element. (88)

Two more passages are relevant here: "The person who says owl-house does not expect his hearer to interpret this as a house that owls fall on or the house my owl flew by. Houses are not generally characterized

---

[15] Since any attribute can in principle be regarded as specific, i.e. as distinguishing the category of entities which possess it from the category of entities which do not possess it (e.g. all those cats crossing a street at the present moment as opposed to all those not doing so) "specific" in Brekle's discussion should presumably be interpreted as meaning something like 'criterial for the definition of some non-ad-hoc category'.

by a regular tendency to be fallen upon by owls. ...the elements of a compound are felt to form a unit that is somehow integral, generic, or necessary, not one that is shortlived or capricious." (91-2)[16] And further: "Notice that compounds are exactly equivalent to all other nouns in the sense that they are in a definitional relation with generic relative clauses. ...Eskimo Dog and Husky are synonyms (neither is a definition of the other), that is, both are defined by the same generic relative [dog of no particular breed that is used by Eskimos]" (97).

The relations that obtain between the constituents of a compound are characterized, as we have seen, by such terms as "habitual", "specific", "generic", "integral". It is certainly true that these and similar criteria do apply to large numbers of compounds, but I think it can be shown that none of them is necessary for compound formation. When I say: "I can't find my bus money," I may perfectly well be referring to some coins I put in a pocket with the intention of using them to pay my bus fare. It is not necessary that I habitually put money in my pocket for this purpose if I want to use the compound bus money. Or take the abovementioned pumpkin bus -- the relation between bus and pumpkin is quite fortuitous, nevertheless the compound seems, under the circumstances, appropriate. The following armchair experiment may be of some help here: Let us imagine two speakers of ENGLISH walking along together. One of them sees a cloud which to him looks somewhat like a kangaroo. He thereupon turns to his companion and says, pointing to the cloud,

(1) Hey, there's a cloud that looks like a kangaroo!
(2) Hey, look at that kangaroo cloud!

(2) seems distinctly odd in this context. The reason for this oddness, we might surmise, is that the relation between cloud and kangaroo is neither habitual or generic. But let us now look at a second situation. Another two people are walking along under the same circumstances, only they belong to a peculiar sect for the adherents of which certain accidental resemblances of real world objects have great importance as supposedly presaging future events. Seeing marsupial-like objects happens, for these people, to be an omen of great prosperity. Under these conditions I submit that (2), as well as (1), would be an appropriate utterance for one of these men. If that is true we must ask what it could be that distinguishes the appropriateness of kangaroo cloud in the second situation from its inappropriateness in the first, given that the referent of the compound is the same in both cases. What seems to be involved is whether a relation is "appropriately classificatory", and while there is clearly wide agreement among the members of a cultural group and

[16] See also footnotes 4 and 9 above.

even among all human beings about the classificatory usefulness of various relations (e.g. the look-like relation, the being-an-instrument-for relation, the being-a-part-of relation), there is room for individual variation. Thus in our culture clouds are not usually classified in terms of their resemblance to animals, but notice that they could be so classified, for instance in a children's game (in which context, as in the one previously mentioned, compounds such as kangaroo cloud, elephant cloud, cat cloud would be quite appropriate.) In general transitory locations in time and space of unique objects are not classificatory features; thus I do not refer to a cat that happens to be sitting in a tree at the time that I observe it as a tree cat, because the property of sitting in a particular tree at a particular time is not relevant to my categorization of cats. The dimension of classificatory relevance that I am trying to define here has something to do with the distinction between naming and description. Anything at all can be described, but only relevant categories are given names (I am talking here about common rather than proper names). To come back to the observers of the kangaroo-shaped cloud: it would be most peculiar if the ENGLISH spoken by the first pair had a noun blick in it which meant 'kangaroo-shaped object', while the existence of such a term in the ENGLISH of the second pair would be no particular cause for surprise.[17] Now compounding is apparently in all languages in which it occurs primarily a device for creating new names, although it should be observed that not all nameworthy referents in such languages for which no name exists will necessarily be designated by a compound. But there is apparently no language, with the possible exception of SANSKRIT, in which the abovementioned cat would be designated by a compound, just as there is no language in which there are monomorphemic words which mean 'an x which happens to be a particular place at a particular time'.[18]

It is unfortunately not easy to give a clear and generally valid definition of "name" as against "description". Is for instance an expression referring to an object and its material composition (e.g. wool glove; notice the

---

[17] The existence of a term meaning 'kangaroo-shaped cloud' would, it is true, be surprising; that is because its existence would imply the existence of other terms meaning 'kangaroo-shaped lake' 'kangaroo-shaped bush', etc. Compounding permits the creation of ad hoc names.

[18] It is not quite clear to me what exactly the situation in non-literary SANSKRIT was. From talking to a number of experts in the field I do at least get the feeling that speakers of Sanskrit were not likely to refer to a bull which happened to be standing in a field by a compound equivalent to field bull.

secondary + primary stress pattern) a name?[19] Certainly material composition is a widely used classificatory feature; why then are expressions such as wool glove not treated formally -- i.e. in terms of stress -- as compounds in ENGLISH? But while some data remain remarkably recalcitrant, it seems to me that in general the naming function of compounds, which I would claim is based on the potentially classificatory nature of the relation between their constituents, furnishes an important criterion for the appropriateness of compounding in the great majority of cases. It might be instructive in this connection to look at an earlier discussion of the naming function of nominal expressions in general. In the Preface to Lees (1960) we find the following passage: "We cannot get along with any single common noun to refer to a familiar object, but must have at every moment modifiers with which to construct new, more complex names to use for all the specific instances of that object which we encounter and talk about. Thus, we cannot, without extensive ambiguity, refer on every occasion to our favorite beverage by means of the single word coffee; instead we name its individual instances with such phrases as my coffee, that cold cup of coffee you left there, some fresh coffee on the shelf, a new brand of coffee, pretty tasteless coffee, Turkish coffee, etc. There is no known limitation on the number of distinct objects for which we must at some time or other have distinctive names, and clearly no dictionary is large enough to contain them all, for a great many of the names which we employ have never before been uttered." (Lees 1960, xvii-xviii). Lees is quite right, except that he missed -- if I am right, that is -- the crucial distinction between naming and description. Thus, of all the examples he gives only Turkish coffee (assuming it is used in the sense in which it is not synonymous with coffee from Turkey) is a name, the other nominal expressions are descriptions. Notice the oddity of some fresh shelf coffee if it is used in the sense of 'some fresh coffee which happens to be on a particular shelf at some particular time'. Notice further that it would surely be most remarkable if ENGLISH had monomorphemic nouns meaning 'my coffee' or 'that cold cup of coffee you left there' or 'a new brand of coffee', although gluck as a noun referring to Turkish coffee does not seem especially surprising. Some correlation seems to exist, also, between names and appropriate responses to the question 'What kind of X is that?' (e.g. What kind of animal is that? -- A zebra.). Of the nominal expressions given by Lees, only Turkish coffee is really a good answer to the question What kind of

---

[19] For some discussion of the different stress patterns associated with N + N sequences see the Appendix.

coffee is that?[20]

One further point must be dealt with briefly. If one looks at appropriate texts in GERMAN one cannot help being struck by the great profusion of ad hoc abstract compounds, many of which have no compounded counterparts in ENGLISH. Thus the following compounds all appear in Brekle (1970) -- and do not by any means constitute a complete listing --

Zweckrelation	Äquivalenzbeziehung
Verfahrensweisen	Sprachsystem
Gegenstandsbegriff	Notationssystem
Informationsdichte	Grammatikmodell
Satzsinn	Sprachtheorie
Teilgebiet	Modellbegriff
Teilsystem	Symbolinventar
Strukturtyp	Kompositionstyp
Syntaxbegriff	Productivitätsgrad
Inhaltsbereich	

Are these compounds appropriate in terms of the notions developed above? It seems to me that a reasonable case can be made for them as not contravening our criteria in spite of their obvious ad hoc nature. Where we are dealing with abstract nouns some of the criteria which militate against a speaker's regarding certain relationships among concrete nouns as classificatory obviously do not apply; thus, there is nothing corresponding to 'transitory location in space and time' for abstract nouns. Thus it could perhaps be said -- very tentatively, to be sure -- that no abstract compound will ever violate the criterion of classificatory relevance, and that at least with respect to that criterion any abstract compound is acceptable. Whether this is true or not, some language-specific qualifications are obviously necessary. For while for example Tugendbegriff seems to be an acceptable GERMAN compound, virtue concept is a rather peculiar ENGLISH one. In general it is obvious that ENGLISH is much less prone to abstract compound formation than GERMAN, although some do occur in ENGLISH too (thus sentence meaning -- cf. Satzsinn above -- seems possible). It is interesting that the great majority of treatments of compounds have concentrated on concrete compounds; clearly much work remains to be done on abstract ones.

---

[20] Unfortunately not all the expressions which I would like to call names constitute appropriate answers to such questions. Thus bus money is apparently not considered a kind of money, i.e. it would not (unlike francs, Canadian dollars, etc.) constitute an appropriate response to a question such as What kind of money did you take with you this morning? It seems that intended purpose does not define kinds of things. Cf. also Henzen (1947, 2-3): "Das Determinativ=Kompositum antwortet auf die Frage: was für ein? (was für eine Art?), die [syntaktische] Verbindung auf die Frage: welcher? wessen? o.ä."

APPENDIX

No attention has so far been paid to the distinction in ENGLISH between compounds (characterized by a primary + tertiary stress pattern) and what Lees has called 'nominal phrases' (secondary + primary), e.g. -- to use an example from Marchand (1969) -- súmmer-hoùse vs. sûmmer résidence.[21] Lees takes the view that there is no systematic derivational (and, presumably, semantic, although he does not say so) difference between the two (Lees 1960, 180-1). Marchand on the other hand sees a radical difference between them:

"Whereas in GERMAN any combination of substantives automatically becomes a compound [i.e. has primary stress on the first element] while a two stressed sb/sb construction does not exist, the rule in ENGLISH is the two-stressed syntactic group while fore-stress is tied up with special grammatical or semantic conditions. Any substantive may be transposed to the role of determinant (modifier, satellite) in a combination where another substantive is the determinatum (head, nucleus). Cf. court jester, government policy: nothing but change of rank is implied in the adjunct position of court and government. The use of substantives as adjuncts is a grammatical phenomenon, the "meaning" implied is merely syntactic and can often be rendered in syntactic terms by a prepositional group such as '(the policy) of the government', '(the jester) of the court'. A GERMAN compound thus may express semantic relations as well as mere syntactic relations while in ENGLISH the latter function is reserved to a syntactic two-stressed group. This would create the difference between sb/sb compounds based on a permanent lexical relation such as súmmer-hoùse and transpositional sb/sb combinations where a mere syntactic relation is expressed such as súmmer résidence. A summer-house is not merely a house inhabited in summer but a house of a particular style and construction which make it suitable for the warm season only. A summer residence, however, is merely '(someone's) residence in the summer', nothing more. Cf. also the difference between Chrístmas trèe and Chrístmas tráffic. In principle, time and space are considered grammatical concepts in ENGLISH. Determination of a substantive by another substantive denoting time or space lead to a double stressed syntactic group." (Marchand 1969, 25).

Marchand distinguishes between transpositional and semantic derivation (cf. also Marchand 1966). In the former, no semantic element is added. Marchand claims that compounds always involve semantic derivation.

---

[21] Marchand calls the latter type two-stressed and represents the stress pattern as '/'.

but to support this position he has to resort to some rather dubious arguments. Thus he regards compounds with the second element store as the result of transpositional derivation (i.e. Bookstore is derived from store for books, or some similar expression, with no resulting semantic difference). To account for the stress pattern he claims that "the frequent occurrence of a word as second consitituent is apt to give compound character to combinations with such words." (Marchand 1969, 23). And further: "The forestress of such combinations is thus due to implicit contrast: each -man, -shop, -store word is automatically stressed on the first member to distinguish the combination from others of the same series." (Marchand 1969, 24). With this contrast ex machina feature added his claim becomes very hard to refute. Still, why is birthday present stressed on the first element? It can be associated with present for a birthday, and there do not seem to be all that many contrasting expressions around. And what about evening course? If it implicitly contrasts with 'course given during the day', why does Christmas traffic not contrast with non-Christmas traffic? And why do some people who say Chrístmas tráffic say hóliday tràffic?

Furthermore, Marchand's transpositional derivations are not in fact innocent of all semantic specialization. Thus I might refer to a car of mine which I permanently keep in New York as my New York car, but I would not use this expression for a car which somebody is driving for me from Providence to Philadelphia and which is currently in New York. And of course I would not cease to refer to my car which is permanently kept in New York as my New York car because it is currently in Stanford. So things are less simple than they seem.

There is furthermore a great deal of dialect variation which is not compatible with the neat distinction that Marchand proposes. Thus we have cŏttage chéese and cŏttage chèese (Marchand would presumably expect this to be forestressed, although he does not claim that 'semantic derivatives' are never two-stressed), chîcken sóup and chícken sòup, âpple píe and ápple pìe. And to complicate matters further some people who say âpple pie say peach pìe.

Given that there are a lot of idiosyncratic factors involved in the compound vs. nominal phrase distinction, it is probably still true that the relations typically embodied in nominal phrases are of a type rather different from what is found in most compounds. Thus time and space relations do often underlie such phrases, as do relations of materal composition (contrast gŏld ring with góldsmìth). And compounds do seem to have a greater tendency to become idiomatized. However, it would appear that the condition of a relation's being "appropriately classificatory" applies to most nominal phrases as well as to compounds.

BIBLIOGRAPHY

Botha, Rudolf P. 1968. The function of the lexicon in transformational generative grammar. Mouton. The Hague.

Brekle, Herbert E. 1966. Syntaktische Gruppe (Adjektiv + Substantiv) vs. Kompositum im modernen Englisch. Linguistics 23, 5-29.

_____. 1970. Generative Satzsemantik und transformationelle Syntax im System der englischen Nominalkomposition. Wilhelm Fink Verlag. München.

Bradley, Henry. 1906. The making of English. Macmillan. London.

Chomsky, Noam. 1957. Syntactic structures. Mouton. The Hague.

_____. 1965. Aspects of the theory of syntax. M.I.T. Press. Cambridge, Mass.

Dokulil, Miloš. 1964. Zum wechselseitigen Verhältnis zwischen Wortbildung und Syntax. Travaux linguistiques de Prague 1, 215-24.

Fillmore, Charles J. 1968. The case for case. Universals in linguistic theory, ed. by Emmon Bach and Robert T. Harms, 1-88. Holt, Rinehart and Winston. New York.

Gleitman, Lila R, and Henry. 1970. Phrase and paraphrase: some innovative uses of language. W.W. Norton. New York.

Henzen, Walter. 1947. Deutsche Wortbildung. Max Niemeyer. Haale, Saale.

Householder, Fred W. 1962. Review of The grammar of English nominalizations, by Robert B. Lees. Word 18, 326-53.

Lees, Robert B. 1960. The grammar of English nominalizations. IJAL 26, No. 3, Part II (= Publication Twelve of the Indiana University Research Center in Anthropology, Folklore, and Linguistics).

_____. 1966. On a transformational analysis of compounds: a reply to Hans Marchand. Indogermanische Forschungen 71, 1-13.

_____. 1970. Problems in the grammatical analysis of English nominal compounds. Progress in linguistics, ed. by Manfred Bierwisch and Karl E. Heidolph, 174-86. Mouton. The Hague.

Li, Charles N. 1970. Topics in semantics. Draft of University of California at Berkeley Ph. D. dissertation.

Marchand, Hans. 1966. Review of Affixal negation in English and other languages, by Karl E. Zimmer. Lg. 42, 134-42.

_____. 1969. The categories and types of present-day English word-formation: a synchronic-diachronic approach. Second edition. Verlag C.H. Beck. Munich.

Motsch, Wolfgang. 1962. Zur Stellung der "Wortbildung" in einem formalen Sprachmodell. Studia Grammatica 1, 31-50.

Rohrer, Christian. 1966. Review of The grammar of English nominalizations, by Robert B. Lees. Indogermanische Forschungen 71, 161-70.

Schachter, Paul. 1962. Review of The grammar of English nominalizations, by Robert B. Lees. IJAL 28, 134-46.

WORKING PAPERS ON LANGUAGE UNIVERSALS

Number Five

May 1971

Language Universals Project
Committee on Linguistics
Stanford University
Stanford, California